

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 17:14:25 ; Search time 41 Seconds
(without alignments)
34.842 Million cell updates/sec

Title: US-09-641-528B-31040
Perfect score: 44
Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 561297

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	65.9	9	18 AAW43834	Specific human leu
2	28	63.6	14	19 AAY21276	Human semaphorin I
3	28	63.6	25	20 AAY27746	Human secreted pro
4	28	63.6	34	17 AAW03650	Human cytomegalovi
5	27	61.4	12	22 AAE09492	Human C3 alpha-1 p
6	27	61.4	35	22 AAE09502	Human C8 alpha-1 p
7	27	61.4	35	23 ABG09097	Breast specific pol
8	26	59.1	25	17 AAW07322	Peptide sequence p
9	26	59.1	35	22 AAW83451	Human immune/haema

10	26	59.1	40	18 AAW20449	H. pylori transpor
11	26	59.1	49	22 ABG52706	Human liver peptid
12	26	59.1	49	22 ABB37858	Peptide #5364 enco
13	26	59.1	49	22 ABB23118	Protein #5117 enco
14	26	59.1	49	22 AAM58489	Human brain expres
15	26	59.1	49	22 AAW70989	Human bone marrow
16	26	59.1	49	22 AAM18760	Peptide #5194 enco
17	26	59.1	49	22 AAM31265	Peptide #5302 enco
18	26	59.1	49	23 ABG40508	Human peptide enco
19	26	59.1	49	24 ABP80853	N. gonorrhoeae ami
20	25	56.8	29	22 ABG49414	Human liver peptid
21	25	56.8	29	22 ABB29414	Peptide #2065 enco
22	25	56.8	29	22 ABB20002	Protein #2001 enco
23	25	56.8	29	22 AAM67779	Human bone marrow
24	25	56.8	29	22 AAM28080	Peptide #2117 enco
25	25	56.8	50	22 AAU48495	Propionibacterium
26	24	54.5	9	21 AAY79337	Equine infectious
27	24	54.5	10	22 AAB96020	Equine infectious
28	24	54.5	12	21 AAY79335	Equine infectious
29	24	54.5	13	21 AAY79334	Equine infectious
30	24	54.5	14	21 AAY79333	Equine infectious
31	24	54.5	14	22 AAM97510	Human peptide #785
32	24	54.5	15	20 AAY45459	Immunogenic peptid
33	24	54.5	15	21 AAY79336	Equine infectious
34	24	54.5	15	24 ABR30426	Human cancer-relat
35	24	54.5	15	24 ABR30581	Human cancer-relat
36	24	54.5	15	24 ABR30959	Human cancer-relat
37	24	54.5	15	24 ABR31232	Human cancer-relat
38	24	54.5	15	24 ABR31233	Human cancer-relat
39	24	54.5	15	24 ABR31567	Human cancer-relat
40	24	54.5	22	10 AAP93517	Amino acid sequenc
41	24	54.5	25	12 AAR13612	Antigenic peptide
42	24	54.5	26	23 AAU97958	Synthetic erythrop
43	24	54.5	27	21 AAY79345	Equine infectious
44	24	54.5	30	23 AAU97950	Synthetic erythrop
45	24	54.5	30	23 AAU97980	Synthetic erythrop

ALIGNMENTS

RESULT 1
AAW43834
ID AAW43834 standard; peptide; 9 AA.
XX
AC AAW43834;
XX
DT 20-APR-1998 (first entry)
XX
DE Specific human leukocyte antigen binding peptide #38.
XX
DE Human leukocyte antigen; HLA; vaccine; bacterial infection; viral;
KW cytotoxic T-cell; CTL; immunogenic peptide; cancer.
XX
OS Synthetic.
OS Herpes simplex virus type 1.
XX
PN WO9734617-A1.
XX
PD 25-SEP-1997.
XX
PF 21-MAR-1997; 97WO-US04451.
XX
PR 20-MAR-1997; 97US-0821739.
XX
PR 21-MAR-1996; 96US-0013833.
XX
(CYTE-) CYTEL CORP.
XX
PI Celis E, Grey HM, Kubo RT, Sette A;
XX
DR WPI; 1997-489250/45.
XX
PT Specific human leukocyte antigen binding peptide - used in vaccines

PT for the treatment and prevention of e.g. bacterial or viral
PT infection and cancer

XX Claim 19; Page 38; 49pp; English.

XX The present sequence represents a specific example of an immunogenic
XX peptide which was used in a new method of inducing a cytotoxic T cell
XX (CTL) response against a preselected antigen in a patient. The method
XX comprises contacting CTLs from the patient with the immunogenic peptide
XX (containing defined motifs) which binds one of the four HLA MHC products
XX HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation
XX constant (Kd) of less than 5 x 10⁻⁷ M. Immunogens are viral, e.g. human
XX immunodeficiency virus type I (HIV-1), hepatitis B virus (HBV) and
XX hepatitis C virus (HCV) or cancer antigens and are used in vaccines for
XX the prevention and treatment of viral infection and cancer. The
XX immunogens may be administered to the patient as a nucleic acid encoding
XX the peptide (gene vaccine).

SQ Sequence 9 AA;

Query Match 65.9%; Score 29; DB 18; Length 9;
Best Local Similarity 71.4%; Pred. No. 9.3e+05;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LERTEVY 9

Db 2 VERTDVI 8

RESULT 2

AAV21276

ID AAY21276 standard; Protein; 14 AA.

XX AAY21276;

22-JUL-1999 (first entry)

Human semaphorin III mutant protein fragment 9.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX frameshift mutation; age-related disease; neurodegenerative disorder;
XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; MSP-A;
XX high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX WO9845322-A2.

PD 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAY75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also

PT for treatment and prevention with specific ribozymes or wild-type
PT RNA

PS Disclosure; Figure 16; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,
XX at an early stage. It is based on the observation that disease may be
XX caused by mutations in RNA rather than DNA. The invention describes the
XX use of neuronal system RNA molecules, specifically proteins including
XX beta-amyloid precursor protein (beta-APP), the microtubule associated
XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX protein-C (HMGP-C) and neuroendocrine specific protein A.

SQ Sequence 14 AA;

Query Match 63.6%; Score 28; DB 19; Length 14;

Best Local Similarity 75.0%; Pred. No. 51;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 TLERTEVY 9

Db 7 TLEGTEVW 14

RESULT 3

AAV27746

ID AAY27746 standard; Protein; 25 AA.

XX AAY27746;

30-JUL-1999 (first entry)

Human secreted protein encoded by gene No. 36.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

XX WO924836-A1.

PD 20-MAY-1999.

XX 04-NOV-1998; 98WO-US23435.

XX 17-NOV-1997; 97US-0066100.

XX 07-NOV-1997; 97US-0064900.

XX 07-NOV-1997; 97US-0064908.

XX 07-NOV-1997; 97US-0064911.

XX 07-NOV-1997; 97US-0064912.

XX 07-NOV-1997; 97US-0064983.

XX 07-NOV-1997; 97US-0064984.

XX 07-NOV-1997; 97US-0064985.

XX 07-NOV-1997; 97US-0064987.

XX 07-NOV-1997; 97US-0064988.

XX 17-NOV-1997; 97US-0066090.

CC The invention relates to a method for detecting Goodpasture antibodies from a bodily fluid or tissue of a patient. The method comprises contacting the bodily fluid or tissue with alpha-3 chain type (IV) collagen polypeptide that contains a conformational epitope for the Goodpasture antibodies. Alpha-3 chain of type (IV) collagen is useful for detecting Goodpasture antibodies from a bodily fluid or tissue from a patient, and for treating Goodpasture syndrome in a patient. The present sequence is human alpha chain peptide used for constructing human alpha1/alpha3 (IV) NC1 fusion protein.

XX Sequence 12 AA;

Query Match 61.4%; Score 27; DB 22; Length 12;

Best Local Similarity 50.0%; Pred. No. 69;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEVY 9
|:|:|:|:
Db 1 TIERSEMF 8

RESULT 6

AAE09502
ID AAE09502 standard; peptide; 35 AA.

XX AAE09502;

DT 19-NOV-2001 (first entry)

DE Human C8 alpha-1 peptide to construct alpha1/alpha3(IV)NC1 protein.

XX Human; alpha-3 chain; type IV collagen; immunosuppressive; therapy;

KW Goodpasture syndrome; C8 alpha-1 peptide.

XX Homo sapiens.

XX US6277558-B1.

XX 21-AUG-2001.

XX 12-NOV-1999; 99US-0439897.

XX 30-NOV-1990; 90US-0621091.

XX 07-MAR-1995; 95US-0399889.

XX 07-OCT-1998; 98US-0167364.

XX (UNIV) UNIV KANSAS MEDICAL CENT.

XX Hudson BG;

XX WPI; 2001-540401/60.

XX Use of alpha (IV) noncollagenous 1 collagen polypeptide for detecting Goodpasture antibodies from bodily fluid/tissue from patient or for treating Goodpasture syndrome by contacting bodily fluid or tissue with the polypeptide -

XX Example 19; Fig 12; 46pp; English.

XX The invention relates to a method for detecting Goodpasture antibodies from a bodily fluid or tissue of a patient. The method comprises contacting the bodily fluid or tissue with alpha-3 chain type (IV) collagen polypeptide that contains a conformational epitope for the Goodpasture antibodies. Alpha-3 chain of type (IV) collagen is useful for detecting Goodpasture antibodies from a bodily fluid or tissue from a patient, and for treating Goodpasture syndrome in a patient. The present sequence is human alpha chain peptide used for constructing human alpha1/alpha3 (IV) NC1 fusion protein.

XX Sequence 35 AA;

Query Match 61.4%; Score 27; DB 22; Length 35;

Best Local Similarity 50.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEVY 9
|:|:|:|:
Db 1 TIERSEMF 8

RESULT 7

ABG90997
ID ABG90997 standard; Protein; 35 AA.

XX ABG90997;

XX 29-NOV-2002 (first entry)

DE Breast specific polypeptide #26.

KW Breast specific; cancer staging; cancer imaging; breast cancer;
KW non-cancerous disease states in breast tissue; gene therapy; vaccine;
KW transgenic animal.

XX Homo sapiens.

XX WO200264741-A2.

XX 22-AUG-2002.

XX 12-FEB-2002; 2002WO-US04134.

XX 13-FEB-2001; 2001US-268289P.

XX (DIAD-) DIADEXUS INC.

XX Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;

XX Liu C, Sun Y;

XX WPI; 2002-657590/70.

XX New breast specific nucleic acids and proteins, for identifying, diagnosing, monitoring, staging, imaging, and treating breast cancer and non-cancerous disease states in breast tissue, or for gene therapy

XX Claim 11; Page 210; 227pp; English.

XX The invention describes a breast specific nucleic acid. The nucleic acids, polypeptides, antibodies, agonists and antagonists, and compositions comprising them are useful for identifying, diagnosing, monitoring, staging, imaging, and treating breast cancer and non-cancerous disease states in breast tissue, in gene therapy, vaccine development, in producing transgenic animals and cells, and in producing engineered breast tissue for treatment and research. The nucleic acids may be used as molecular markers for detecting breast cancer for accurate staging of the disease and monitoring the progress of cancer treatments, CC and as hybridisation probes or primers. This is the amino acid sequence of a breast specific protein of the invention.

XX Sequence 35 AA;

Query Match 61.4%; Score 27; DB 23; Length 35;

Best Local Similarity 71.4%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLERTEV 8
|:|:|:|:
Db 22 TLRSTEI 28

RESULT 8

AAW07322
ID AAW07322 standard; peptide; 25 AA.

XX AAW07322;

XX 25-MAR-2003 (updated)
 DT 11-FEB-1997 (first entry)
 XX
 XX Peptide sequence present in pro-RIP1 but not in mature protein.
 XX
 XX Maize; ribosomal inhibitory protein; leaf; clone; probe; immunotoxin;
 KW cross-hybridisation; open reading frame; agricultural pesticide;
 KW antiviral agent; protein translation; plant; eukaryotic; ribosome.
 XX
 OS Synthetic.
 XX
 PN US552140-A.
 XX
 XX 03-SEP-1996.
 PD
 XX 25-JUL-1994; 94US-0279996.
 PF
 XX 08-SEP-1992; 92US-0941651.
 PR 25-JUL-1994; 94US-0279996.
 XX
 PA (UYNC-) UNIV NORTH CAROLINA STATE.
 XX
 XX Bass HW, Boston RS, Obrian GR;
 XX
 XX WPI; 1996-412013/41.
 DR N-PSDB; AAT42304.
 XX
 XX Maize ribosome-inactivating protein - useful as agricultural
 PT pesticide, immunotoxin or antiviral agent, and to arrest protein
 PT translation in biochemical assays
 XX
 PS Disclosure; Column 21-22; 21pp; English.
 XX
 CC This sequence corresp. to amino acids 193-217 of the maize ribosomal
 CC inhibitory protein 1 (RIP1). The sequence was used to generate a probe
 CC which does not bind to the RIP2 gene sequence (AAT42302). The amino
 CC acid sequence is present in the pro-RIP1 form and absent in the mature
 CC protein. The RIP2 gene can be used as an agricultural pesticide, an
 CC immunotoxin or antiviral agent or to arrest protein translation in
 CC biochemical assays. RIPs acts mainly against non-plant eukaryotic
 CC ribosomes by enzymatically attacking the large 60S subunit and
 CC irreversibly modifying the large ribosomal RNA.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 25 AA;
 Query Match 59.1%; Score 26; DB 17; Length 25;
 Best Local Similarity 75.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATLETEV 8
 Db |||||
 2 ATLEEEV 9
 RESULT 9
 AAM83451
 ID AAM83451 standard; Protein; 35 AA.
 XX
 AC AAM83451;
 XX
 XX 07-NOV-2001 (first entry)
 DE
 DE Human immune/haematopoietic antigen SEQ ID NO:11044.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX

PD 09-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX
 XX
 XX

DR N-PSDB; AAK56232.
 XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX Claim 11; SEQ ID NO 11044; 3071pp + Sequence Listing; English.
 PS
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX Sequence 35 AA;
 SQ
 Query Match 59.1%; Score 26; DB 22; Length 35;
 Best Local Similarity 57.1%; Pred. No. 3.5e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 LERTEVY 9
 Db 12 LERAQIY 18
 RESULT 10
 AAW20449
 ID AAW20449 standard; Protein; 40 AA.
 XX
 AC AAW20449;
 XX
 DT 11-JUL-1997 (first entry)
 XX
 DE H. pylori transporter protein, 35428912.aa.
 XX
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacterium; life cycle; activator;
 KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
 KW diagnosis.
 XX
 OS Helicobacter pylori.
 XX
 PN WO9640893-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-US09122.
 XX
 PR 01-APR-1996; 96US-0630405.
 PR 07-JUN-1995; 95US-0487032.
 XX
 PA (ASTR) ASTRA AB.
 XX
 PI Berglinth OT, Smith D, Mellgaard BL;
 XX
 DR WPI; 1997-052306/05.
 DR N-PSDB; AAT67621.
 XX
 XX Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter

XX Claim 56; Page 622; 1481pp; English.

XX This sequence represents a H. pylori transporter protein.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds,

CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from

CC overlapping contigs generated by mechanically shearing the bacterial

CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,

CC and the predicted coding regions defined by computer evaluation. To

CC identify likely H. pylori antigens for vaccine development, the amino

CC acid sequences predicted from various ORF were analysed for significant

CC homology to other known or exported membrane proteins. Having identified

CC and determined the sequences of interest, particular regions can be

CC isolated from H. pylori by PCR amplification for recombinant polypeptide

CC production, e.g. in E. coli hosts.

XX Sequence 40 AA;

SQ

Query Match 59.1%; Score 26; DB 18; Length 40;

Best Local Similarity 62.5%; Pred. No. 4e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLTEREV 8

DB :|||

21 STIERCEV 28

RESULT 11

ABG52706

ID ABG52706 standard; Peptide; 49 AA.

XX AC ABG52706;

XX DT 25-FEB-2003 (first entry)

XX DE Human liver peptide, SEQ ID No 31354.

XX DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

XX OS Homo sapiens.

XX PN WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000664.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analysing gene expression in human adult liver -

XX Claim 27; SEQ ID No 31354; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult

CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes

CC identified may be involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which

CC is associated with coronary heart disease. ABG47348-ABG59930 represent

CC human liver single exon encoded peptides of the invention.

CC Note: The sequence information for this patent does not appear in the

CC printed specification but was obtained in electronic format directly

CC from WIPO at fp.wipo.int/pub/published_pct_sequences.

XX Sequence 49 AA;

SQ

Query Match 59.1%; Score 26; DB 22; Length 49;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RTEVY 9

DB :|||

1 RTEVY 5

RESULT 12

ABB37858

ID ABB37858 standard; Peptide; 49 AA.

XX AC ABB37858;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #5364 encoded by human foetal liver single exon probe.

XX DE Human; foetal liver; gene expression; single exon nucleic acid probe.

KW Homo sapiens.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analysing gene expression in human foetal liver -

XX Claim 27; SEQ ID No 30493; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at fp.wipo.int/pub/published_pct_sequences.

XX Sequence 49 AA;

SQ

Query Match 59.1%; Score 26; DB 22; Length 49;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RTEVY 9
DB 1 RTEVY 5

RESULT 13

ABB23118
ID ABB23118 standard; Protein; 49 AA.

XX AC ABB23118;
XX DT 23-JAN-2002 (first entry)

DE DE Protein #5117 encoded by probe for measuring heart cell gene expression.

XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.

XX OS Homo sapiens.

XX XX WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488899/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human hearts -

XX PS Claim 15; SEQ ID No 24888; 530pp; English.

XX CC The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABA21535-ABA41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 49 AA;

Query Match 59.1%; Score 26; DB 22; Length 49;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RTEVY 9
DB 1 RTEVY 5

RESULT 14

AAM58489
ID AAM58489 standard; Protein; 49 AA.

XX AC AAM58489;
XX DT 05-NOV-2001 (first entry)

DE DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30594.

XX KW Human; brain expressed exon; gene expression analysis; probe;
XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX KW epilepsy; cancer.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human brains -

XX PS Example 4; SEQ ID NO: 30594; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.

XX SQ Sequence 49 AA;

Query Match 59.1%; Score 26; DB 22; Length 49;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RTEVY 9
DB 1 RTEVY 5

RESULT 15

AM70989
ID AM70989 standard; Protein; 49 AA.

XX AC AM70989;
XX DT 06-NOV-2001 (first entry)

DE DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31295.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.


```
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX XX
XX PD 09-AUG-2001.
XX XX
XX PF 30-JAN-2001; 2001WO-US00668.
XX XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX
XX DR WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human bone marrow -
XX PS Example 4; SEQ ID NO: 31295; 658pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention.
XX SQ Sequence 49 AA;
XX
XX Query Match 59.1%; Score 26; DB 22; Length 49;
XX Best Local Similarity 100.0%; Pred. No. 5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 RTEVY 9
DB 1 RTEVY 5
```

Search completed: January 2, 2004, 17:33:51
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 17:33:06 ; Search time 20 Seconds
(without alignments)
43.276 Million cell updates/sec

Title: US-09-641-528B-31040

Perfect score: 44

Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 11832

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	56.8	44	2	H64308
2	24	54.5	21	2	A38837
3	24	54.5	32	2	A32502
4	24	54.5	33	2	PH1741
5	24	54.5	37	2	B95215
6	24	54.5	40	2	H95183
7	24	54.5	46	2	S26260
8	23	52.3	27	2	A28391
9	23	52.3	30	2	S14062
10	23	52.3	30	2	C82092
11	23	52.3	33	2	A81046
12	23	52.3	38	2	C81501
13	23	52.3	44	2	JH0350
14	22	50.0	18	2	D49570
15	22	50.0	22	2	S62670
16	22	50.0	24	2	S47563
17	22	50.0	34	2	E82458
18	22	50.0	34	2	S77646
19	22	50.0	37	2	AH0844
20	22	50.0	38	1	F2RZL
21	22	50.0	38	1	F2WTL
22	22	50.0	38	1	F2SKL
23	22	50.0	38	2	T07254
24	22	50.0	38	2	S78339
25	22	50.0	38	2	S04064
26	22	50.0	38	2	S05685
27	22	50.0	38	2	S31821
28	22	50.0	38	2	A05048
29	22	50.0	38	2	S03193

30 22 50.0 38 2 S00691 photosystem II pro
31 22 50.0 38 2 S58566 photosystem II pro
32 22 50.0 38 2 C48310 photosystem II pro
33 22 50.0 38 2 S51366 photosystem II pro
34 22 50.0 38 2 S28056 photosystem II pro
35 22 50.0 38 2 T07477 photosystem II pro
36 22 50.0 39 2 S75180 photosystem II pro
37 22 50.0 39 2 AH2286 photosystem II pro
38 22 50.0 42 2 E46104 polymerase L - Mok
39 22 50.0 44 2 A82951 50S ribosomal prot
40 22 50.0 44 2 S26879 photosystem II pro
41 22 50.0 45 2 H44530 T-cell receptor al
42 22 50.0 47 2 AF3178 hypothetical prote
43 22 50.0 49 1 TOB03 thymoposetin III -
44 22 50.0 50 2 S41323 8K protein - potat
45 21 47.7 22 2 I41299 T-cell receptor be

ALIGNMENTS

RESULT 1

H64308

hypothetical protein MJ0072 - Methanococcus jannaschii

C;Species: Methanococcus jannaschii

C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C;Accession: H64308

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.;
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hursb, M.A.

Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A;Reference number: A64300; MUID:96337999; PMID:8688087

A;Accession: H64308

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-44 <BUL>

A;Cross-references: GB:U67464; GB:L77117; NID:gi590852; PIDN:AAB98054.1; PID:gi590858; T

C;Genetics:

A;Map position: FOR67562-67696

Query Match 56.8%; Score 25; DB 2; Length 44;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ERTEVY 9

Db 30 ERAEIV 35

RESULT 2

A38837

T-cell receptor beta chain precursor V region (12.6) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 23-Jul-1999

C;Accession: A38837

R;Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayi, G.;

Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991

A;Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumatoid

A;Reference number: A41299; MUID:92020887; PMID:1656449

A;Accession: A38837

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-21 <UEM>

A;Cross-references: GB:S57596; NID:g236336; PIDN:AAB19965.1; PID:g236337

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: T-cell receptor

Query Match 54.5%; Score 24; DB 2; Length 21;

Best Local Similarity 55.6%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
||:|
Db 9 ATRSQTSVY 17

RESULT 3

A32502
T-cell receptor delta chain CDN26, thymus - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 05-Oct-1989 #sequence_revision 05-Oct-1989 #text_change 30-May-1997
C:Accession: A32502
R:Jacy, M.J.; McNeil, L.K.; Roth, M.E.; Kranz, D.M.
Proc. Natl. Acad. Sci. U.S.A. 86, 1023-1026, 1989
A:Title: T-cell receptor delta-chain diversity in peripheral lymphocytes.
A:Reference number: A31461; MUID:99128840; PMID:2783779
A:Accession: A32502
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-32 <LAC>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 54.5%; Score 24; DB 2; Length 32;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
||:|
Db 6 ALMERABGY 14

RESULT 4

PH1741
Ig heavy chain V region (clone NP-12-7) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1741
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1741
A:Molecule type: mRNA
A:Residues: 1-33 <MCH>
A:Experimental source: B cell
A>Note: the authors translated the codon ATA for residue 24 as Phe
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 54.5%; Score 24; DB 2; Length 33;
Best Local Similarity 37.5%; Pred. No. 1.8e+02;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
::|
Db 18 SIQSTEVY 25

RESULT 5

B95215
hypothetical protein SP1842 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: B95215
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: B95215
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-37 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75915.1; PID:gl4973344; GSPDB:GN00164; TIGR:SP4;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI842

Query Match 54.5%; Score 24; DB 2; Length 37;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
||:|
Db 19 LEKVELY 25

RESULT 6

H95183
hypothetical protein SPI579 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95183
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95183
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-40 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75665.1; PID:gl4973070; GSPDB:GN00164; TIGR:SP4;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI579

Query Match 54.5%; Score 24; DB 2; Length 40;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERTEVY 9
||:|
Db 26 EKTEVF 31

RESULT 7

S26260
T-cell receptor beta chain (Vbeta 12.3) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S26260
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
J. Immunol. 147, 4360-4365, 1991
A:Title: New human V-beta genes and polymorphic variants.
A:Reference number: S26255; MUID:92091749; PMID:1684376
A:Accession: S26260
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-46 <PLA>
A:Cross-references: EMBL:X57609; NID:g23181; PIDN:CAA40835.1; PID:g930037
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 54.5%; Score 24; DB 2; Length 46;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
||:|
Db 33 ATRSQTSVY 41

RESULT 8
A28391
hypothetical protein 1 - Streptomyces albidoflavus (fragment)
C;Species: Streptomyces albidoflavus
C;Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 11-Jan-2000
C;Accession: A28391
R;Long, C.M.; Virolle, M.J.; Chang, S.Y.; Chang, S.; Bibb, M.J.
J. Bacteriol. 169, 5745-5754, 1987
A;Title: Alpha-amylase gene of Streptomyces limosus: nucleotide sequence, expression map
A;Reference number: A91856; MUID:98058792; PMID:3500166
A;Accession: A28391
A;Molecule type: DNA
A;Residues: 1-27 <LON>
A;Note: the source is designated as Streptomyces limosus
C;Superfamily: lac repressor

Query Match 52.3%; Score 23; DB 2; Length 27;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTE 7
Db 2 TVQRTE 7

RESULT 9
S14062
hypothetical protein 1 - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 19-Mar-1997 #sequence_revision 26-Feb-1998 #text_change 11-Jan-2000
C;Accession: S14062
R;Vigal, T.; Gil, J.A.; Daza, A.; Garcia-Gonzalez, M.D.; Martin, J.F.
Mol. Gen. Genet. 225, 278-288, 1991
A;Title: Cloning, characterization and expression of an alpha-amylase gene from Streptomyces
A;Reference number: S14062; MUID:9117128; PMID:1900915
A;Accession: S14062
A;Molecule type: DNA
A;Residues: 1-30 <VIG>
A;Experimental source: strain IMRU3570
C;Superfamily: lac repressor

Query Match 52.3%; Score 23; DB 2; Length 30;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLERTE 7
Db 5 TVQRTE 10

RESULT 10
C82092
hypothetical protein VC2306 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: C82092
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-30 <HEI>
A;Cross-references: GB:AE004301; GB:AE003852; NID:99656865; PIDN:AAF95450.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2306
A;Map position: 1

Query Match 52.3%; Score 23; DB 2; Length 30;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ERTEVY 9
Db 24 ERADVY 29

RESULT 11

A81046
cryptic plasmid protein A-related protein NMB1754 [imported] - Neisseria meningitidis (serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: A81046
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Zi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: A81046
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-33 <TET>
A;Cross-references: GB:AE002525; GB:AE002098; NID:G7227004; PIDN:AAF42095.1; PID:G722700
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1754

Query Match 52.3%; Score 23; DB 2; Length 33;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTEV 8
Db 9 LKRTET 14

RESULT 12

C81501
hypothetical protein CP1128 [imported] - Chlamydia pneumoniae (strain AR39)
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: C81501
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81501
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-38 <REA>
A;Cross-references: GB:AE002268; GB:AE002161; NID:G7190029; PIDN:AAF38894.1; PID:G719004
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP1128

Query Match 52.3%; Score 23; DB 2; Length 38;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERTEVY 9
Db 7 EKTEIF 12

RESULT 13

JH0350
T-cell receptor beta chain V region (P1F12C4) - mouse (fragment)
C;Species: Mus musculus (house mouse)

C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 23-Jul-1999
 C;Accession: JH0350
 R;Taylor, A.H.; Haberman, A.M.; Gerhard, W.; Caton, A.J.
 J. Exp. Med. 172, 1643-1651, 1990
 A;Title: Structure-function relationships among highly diverse T cells that recognize a
 A;Reference number: JH0333; MUID:91079767; PMID:1701821
 A;Accession: JH0350
 A;Molecule type: mRNA
 A;Residues: 1-44 <TAY>
 A;Cross-references: GB:M34211; NID:G201426; PIDN:AAA58766.1; PID:G201427
 C;Comment: This T-cell receptor recognizes a determinant from influenza virus hemagglutinin
 C;Superfamily: immunoglobulin V region; immunoglobulin homology
 C;Keywords: receptor; T-cell

Query Match 52.3%; Score 23; DB 2; Length 44;
 Best Local Similarity 55.6%; Pred. No. 4e+02; 3; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
 || : ||
 Db 13 ATPSQTSVY 21

RESULT 14
 D49570
 plasma membrane calcium pump - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 12-Dec-1997
 C;Accession: D49570
 R;Stauffer, T.P.; Hilfiker, H.; Carafoli, E.; Strehler, E.E.
 J. Biol. Chem. 268, 25993-26003, 1993
 A;Title: Quantitative analysis of alternative splicing options of human plasma membrane
 A;Reference number: A49570; MUID:94064681; PMID:8245032
 A;Accession: D49570
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-18 <STA>
 A;Experimental source: cerebral cortex
 A;Note: sequence inconsistent with nucleotide translation
 A;Note: sequence extracted from NCBI backbone (NCBIN:139934, NCBIPI:139939)
 C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C;Keywords: ATP

Query Match 50.0%; Score 22; DB 2; Length 18;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LERTEV 8
 || || ||
 Db 12 LETTEV 17

RESULT 15
 S62670
 folate binding protein II - rat (fragment)
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
 C;Accession: S62670
 R;da Costa, M.; Rothenberg, S.P.
 Biochim. Biophys. Acta 1292, 23-30, 1996
 A;Title: Purification and characterization of folate binding proteins from rat placenta.
 A;Reference number: S62669; MUID:96139323; PMID:8547345
 A;Accession: S62670
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-22 <DAC>

Query Match 50.0%; Score 22; DB 2; Length 22;
 Best Local Similarity 62.5%; Pred. No. 3.1e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATLERTEV 8
 || || ||:

Db 2 ATRARTEL 9

Search completed: January 2, 2004, 17:36:11
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 17:28:25 ; Search time 11 Seconds
(without alignments)
38.476 Million cell updates/sec

Title: US-09-641-528B-31040
Perfect score: 44
Sequence: 1 ATLETEVY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 4018

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	56.8	44	1 Y072 METJA	Q60378 methanococ
2	23	52.3	33	1 PBAN_LYMDI	P43511 lymantria d
3	22	50.0	14	1 PH1_PRUSE	P29263 prunus sero
4	22	50.0	15	1 PH3_PRUSE	P29265 prunus sero
5	22	50.0	16	1 PH2_PRUSE	P29264 prunus sero
6	22	50.0	37	1 PSBL_ARATH	P29301 arabidopsis
7	22	50.0	37	1 PSBL_ORYSA	P12166 oryza sativ
8	22	50.0	38	1 PSBL_CHLEU	P46306 chlamydomon
9	22	50.0	38	1 PSBL_CHLRE	P32974 chlamydomon
10	22	50.0	38	1 PSBL_CHLVU	P56339 chlorella v
11	22	50.0	38	1 PSBL_EUGGR	P12228 euglena gra
12	22	50.0	38	1 PSBL_MARPO	P12165 marchantia
13	22	50.0	38	1 PSBL_MESVI	Q2muq2 mesostigma
14	22	50.0	38	1 PSBL_ODOSI	P49514 odontella s
15	22	50.0	38	1 PSBL_PINTH	P16171 pinus thunb
16	22	50.0	39	1 PSBL_PSYN3	Q55354 synchocyst
17	22	50.0	44	1 RL34_PSEAE	P29436 pseudomonas
18	22	50.0	49	1 THPS_BOVIN	P01251 bos taurus
19	21	47.7	26	1 YFHA_KLEPN	P21710 klebsiella
20	21	47.7	39	1 PSBF_PSINU	Q8w105 psilotum n
21	21	47.7	42	1 Y916_TREPA	Q83866 treponema p
22	21	47.7	49	1 R331_LISMO	Q92c19 listeria mo
23	20	45.5	17	1 A1YS_MYCPH	P81528 mycobacteri
24	20	45.5	21	1 MCT3_MOUSE	P21843 mus musculu
25	20	45.5	24	1 OS94_RAT	P83581 rattus norv
26	20	45.5	33	1 GAST_DIDMA	P33713 didelphis m
27	20	45.5	41	1 TYB9_BOVIN	P21752 bos taurus
28	20	45.5	41	1 TYB9_PIG	P21753 sus scrofa
29	20	45.5	43	1 TYB0_HUMAN	P13472 homo sapien
30	20	45.5	43	1 TYB4_XENLA	P18758 xenopus lae
31	20	45.5	49	1 R332_BACLI	P35870 bacillus li
32	19	43.2	18	1 AL13_CARMA	P81816 carcinus ma
33	19	43.2	31	1 PSAM_SYNEL	P25903 synchococc

ALIGNMENTS

RESULT 1

Y072_METJA 19 43.2 32 1 CAL_ANGUA
ID Y072_METJA STANDARD; PRT; 44 AA.
AC Q60378;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0072.
GN MJ0072.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67464; AAB98054.1; -;
DR PIR; H64308; H64308.
DR TIGR; MJ0072; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 44 AA; 4889 MW; C742DB9EBBC1F451 CRC64;

Query Match 56.8%; Score 25; DB 1; Length 44;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ERTEVY 9
|||
Db 30 ERAEIV 35

RESULT 2

PBAN_LYMDI STANDARD; PRT; 33 AA.
ID PBAN_LYMDI
AC P43511;


```

KW Glycosidase; Hydrolase; Glycoprotein; Multigene family.
FT NON TER 16 16
SQ SEQUENCE 16 AA; 1703 MW; FC4D7F4FB90CFE01 CRC64;

Query Match 50.0%; Score 22; DB 1; Length 16;
Best Local Similarity 83.3%; Pred.No. 92;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATLTR 6
   ||| ||
Db 10 ATLKRT 15

RESULT 6
PSBL ARATH
ID ID PSBL ARATH STANDARD; PRT; 37 AA.
AC P29301;
DT 01-DSC-1992 (Rel. 24, Created)
DT 15-DSC-1998 (Rel. 37, Last sequence update)
DT 15-SBP-2003 (Rel. 42, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
PSBL OR ATCG00560.
OS Arabidopsis thaliana (Mouse-ear cress),
OS Antirrhinum majus (Garden snapdragon), and
OS Lotus japonicus.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702, 4151, 34305;
RN [1]
SEQUENCE FROM N.A.
RP SPECIES=A.thaliana; STRAIN=cv. Columbia;
RC MEDLINE=20039611; PubMed=10574454;
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;
RT "Complete structure of the chloroplast genome of Arabidopsis
thaliana.";
RL DNA Res. 6:283-290(1999).
RN [2]
SEQUENCE FROM N.A.
RP SPECIES=A.majus; STRAIN=cv. Sippe 50; TISSUE=Leaf;
RC MEDLINE=92191997; PubMed=1547774;
RA Kudla J., Igloi G.L., Metzlaiff M., Hagemann H., Koesel H.;
RT "RNA editing in tobacco chloroplasts leads to the formation of a
translatable psbL mRNA by a C to U substitution within the initiation
codon.";
RL EMBO J. 11:1099-1103(1992).
RN [3]
SEQUENCE FROM N.A.
RP SPECIES=L.japonicus; STRAIN=Accession MG-20;
RC MEDLINE=21082929; PubMed=11214967;
RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;
RT "Complete structure of the chloroplast genome of a legume, Lotus
japonicus.";
RL DNA Res. 7:323-330(2000).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AP000423; BAA84400.1; -.
CC EMBL; X63206; CAA44889.1; -.
CC EMBL; AP002983; BAB33211.1; -.
CC EMBL; S19786; F2SK.
CC InterPro; IPR003372; PSII_Psbl.
CC Pfam; PF02419; Psbl; 1.

```

DR ProDom; PD005113; PSII_PeblL; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
FT INIT MET 0 0 BY SIMILARITY.
SQ SEQUENCE 37 AA; 4339 MW; CC537AEC4C630A84 CRC64;

Query Match 50.0%; Score 22; DB 1; Length 37;
Best Local Similarity 57.1%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LERTEVY 9
DB 12 LARTSLY 18
| | | | |
| | | | |

RESULT 7
PSBL ORYSA STANDARD; PRT; 37 AA.
ID PSBL ORYSA STANDARD; PRT; 37 AA.
AC P1216; O47030; P12167; Q34007;
DT 01-DEC-1989 (Rel. 12, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Oryza sativa (Rice),
OS Nicotiana tabacum (Common tobacco),
OS Hordeum vulgare (Barley),
OS Triticum aestivum (Wheat),
OS Secale cereale (Rye),
OS Zea mays (Maize),
OS Pisum sativum (Garden pea),
OS Spinacia oleracea (Spinach),
OS Capsicum annuum (Bell pepper),
OS Mesembryanthemum crystallinum (Common ice plant),
OS Beta vulgaris (Sugar beet), and
OS Populus deltoides (Poplar).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
NCBI TaxID=4530, 4097, 4513, 4565, 4550, 4577, 3888, 3562, 4072, 3544,
OX 161934, 3696;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=N.tabacum; STRAIN=cv. Nipponbare;
RX MEDLINE=89364698; PubMed=2770692;
RA Hiraetsuka J., Yoshida H., Whittier R., Ishibashi T., Sakamoto M.,
RA Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,
RA Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Sugura M.;
RA "The complete sequence of the rice (Oryza sativa) chloroplast genome:
RA intermolecular recombination between distinct trRNA genes accounts for
RA a major plastid DNA inversion during the evolution of the cereals.";
RL Mol. Gen. Genet. 217:185-194(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=N.tabacum; STRAIN=cv. Bright Yellow 4;
RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,
RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,
RA Yamauchi-Shinozaki K., Ohto C., Torazawa K., Meng B.-Y., Sugita M.,
RA Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,
RA Tohdoh N., Shimada H., Sugura M.;
RA "The complete nucleotide sequence of the tobacco chloroplast genome:
RA its gene organization and expression.";
RL EMBO J. 5:2043-2049(1986).
RN [3]
RP RNA EDITING OF INITIATOR CODON.
RC SPECIES=N.tabacum;
RX MEDLINE=97076156; PubMed=8918473;
RA Chaudhuri S., Maliga P.;
RA "Sequences directing C to U editing of the plastid psbL mRNA are
RA located within a 22 nucleotide segment spanning the editing site.";
RL EMBO J. 15:5958-5964(1996).
RN [4]
RP SEQUENCE FROM N.A.

DR ProDom; PD005113; PSII_Pebl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
FT INIT MET 0 0 BY SIMILARITY.
SQ SEQUENCE 37 AA; 4339 MW; CC537AEC4C630A84 CRC64;

Query Match 50.0%; Score 22; DB 1; Length 37;
Best Local Similarity 57.1%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LERTEVY 9
DB 12 LARTSLY 18
| | | | |
| | | | |

RESULT 7
PSBL ORYSA STANDARD; PRT; 37 AA.
ID PSBL ORYSA STANDARD; PRT; 37 AA.
AC P1216; O47030; P12167; Q34007;
DT 01-DEC-1989 (Rel. 12, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Oryza sativa (Rice),
OS Nicotiana tabacum (Common tobacco),
OS Hordeum vulgare (Barley),
OS Triticum aestivum (Wheat),
OS Secale cereale (Rye),
OS Zea mays (Maize),
OS Pisum sativum (Garden pea),
OS Spinacia oleracea (Spinach),
OS Capsicum annuum (Bell pepper),
OS Mesembryanthemum crystallinum (Common ice plant),
OS Beta vulgaris (Sugar beet), and
OS Populus deltoides (Poplar).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI TaxID=4530, 4097, 4513, 4565, 4550, 4577, 3888, 3562, 4072, 3544,
OX 161934, 3696;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=N.tabacum; STRAIN=cv. Nipponbare;
RX MEDLINE=89364698; PubMed=2770692;
RA Hirasaka J., Yoshida H., Whittier R., Ishibashi T., Sakamoto M.,
RA Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,
RA Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Sugita M.;
RA "The complete sequence of the rice (Oryza sativa) chloroplast genome:
RA intermolecular recombination between distinct trRNA genes accounts for
RA a major plastid DNA inversion during the evolution of the cereals.";
RL Mol. Gen. Genet. 217:185-194(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=N.tabacum; STRAIN=cv. Bright Yellow 4;
RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,
RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,
RA Yamauchi-Shinozaki K., Ohto C., Torazawa K., Meng B.-Y., Sugita M.,
RA Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,
RA Tohdoh N., Shimada H., Sugita M.;
RA "The complete nucleotide sequence of the tobacco chloroplast genome:
RA its gene organization and expression.";
RL EMBO J. 5:2043-2049(1986).
RN [3]
RP RNA EDITING OF INITIATOR CODON.
RC SPECIES=N.tabacum;
RX MEDLINE=97076156; PubMed=8918473;
RA Chaudhuri S., Maliga P.;
RA "Sequences directing C to U editing of the plastid psbL mRNA are
RA located within a 22 nucleotide segment spanning the editing site.";
RL EMBO J. 15:5958-5964(1996).
RN [4]
RP SEQUENCE FROM N.A.

RC SPECIES=H.vulgare; STRAIN=cv. Sabarlis;
RX MEDLINE=89240046; PubMed=2654886;
RA Chakhmakheva O.G., Andreeva A.V., Buryakova A.A., Reverdatto S.V.,
RA Efimov V.A.;
RT "Nucleotide sequence of the barley chloroplast psbE, psbF genes and
RT flanking regions.";
RL Nucleic Acids Res. 17:2858-2858(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=H.vulgare; STRAIN=cv. Sabarlis;
RX MEDLINE=92207253; PubMed=1804121;
RA Efimov V.A., Andreeva A.V., Reverdatto S.V., Chakhmakheva O.G.;
RT "Photosystem II of rye. Nucleotide sequence of the psbB, psbC, psbE,
RT psbF, psbH genes of rye and chloroplast DNA regions adjacent to
RT them.";
RL Bioorg. Khim. 17:1369-1385(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=T.aestivum; STRAIN=cv. SENTRY; TISSUE=Leaf;
RA Webber A.N., Hird S.M., Packman L.C., Dyer T.A., Gray J.C.;
RT "A photosystem II polypeptide is encoded by an open reading frame
RT co-transcribed with genes for cytochrome b-559 in wheat chloroplast
RT DNA.";
RL Plant Mol. Biol. 12:141-151(1989).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=T.aestivum; STRAIN=cv. Chinese Spring;
RA Ogihara Y., Isono K., Kojima T., Endo A., Hanaoka M., Shiina T.,
RA Terachi T., Uetani S., Murata M., Mori N., Takumi S., Ikeo K.,
RA Gojobori T., Murai K., Murai K., Matsuoka Y., Ohnishi Y., Tajiri H.,
RA Tanigawa K.;
RT "Chinese spring wheat (Triticum aestivum L.) chloroplast genome:
RT complete sequence and contig clones.";
RL Plant Mol. Biol. Rep. 18:243-253(2000).
RN [8]
RP RNA EDITING OF INITIATOR CODON.
RC SPECIES=T.aestivum;
RX MEDLINE=92191997; PubMed=1547774;
RA Kudla J., Igloi G.L., Metzlauff M., Hagemann R., Koessel H.;
RT "RNA editing in tobacco chloroplasts leads to the formation of a
RT translatable psbL mRNA by a C to U substitution within the initiation
RT codon.";
RL EMBO J. 11:1099-1103(1992).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=S.cereale;
RX MEDLINE=89160331; PubMed=2646599;
RA Zolotarev A.S., Kolosov V.L.;
RT "Nucleotide sequence of the rye chloroplast DNA fragment, comprising
RT psbE and psbF genes.";
RL Nucleic Acids Res. 17:1760-1760(1989).
RN [10]
RP SEQUENCE FROM N.A.
RC SPECIES=S.cereale;
RX MEDLINE=90073796; PubMed=2686655;
RA Kolosov V.L., Klezovich O.N., Abdulaev N.G., Zolotarev A.S.;
RT "Photosystem II of rye. Nucleotide sequence of genes psbE, psbF, psbL
RT and OFC40 of chloroplast DNA.";
RL Bioorg. Khim. 15:1284-1286(1989).
RN [11]
RP SEQUENCE FROM N.A.
RC SPECIES=Maize;
RA Haley J., Bogorad L.;
RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RC SPECIES=Maize;
RX MEDLINE=95395841; PubMed=7666415;
RA Maier R.M., Neckermaier K., Igloi G.L., Koessel H.;
RT "Complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.";
RL J. Mol. Biol. 251:614-628(1995).

RN [13]
RP SEQUENCE FROM N.A.
RC SPECIES=P.esativum;
RX MEDLINE=89354671; PubMed=2766383;
RA Willey D.L., Gray J.C.;
RT "Two small open reading frames are co-transcribed with the pea
RT chloroplast genes for the polypeptides of cytochrome b-559.";
RL Curr. Genet. 15:213-220(1989).
RN [14]
RP SEQUENCE FROM N.A., AND RNA EDITING OF INITIATOR CODON.
RC SPECIES=S.oleracea;
RX MEDLINE=93360903; PubMed=8355656;
RA Bock R., Hagemann R., Koessel H., Kudla J.;
RT "Tissue- and stage-specific modulation of RNA editing of the psbF and
RT psbL transcript from spinach plastids -- a new regulatory mechanism?";
RL Mol. Gen. Genet. 240:238-244(1993).
RN [15]
RP SEQUENCE FROM N.A.
RC SPECIES=S.oleracea; STRAIN=cv. Geant d'hiver, and cv. Monatol;
RX MEDLINE=21187424; PubMed=11292076;
RA Schmitz-Linneweber C., Maier R.M., Alcaraz J.-P., Cottet A.,
RA Herrmann R.G., Mache R.;
RT "The plastid chromosome of spinach (Spinacia oleracea): complete
RT nucleotide sequence and gene organization.";
RL Plant Mol. Biol. 45:307-315(2001).
RN [16]
RP SEQUENCE OF 1-12 FROM N.A.
RC SPECIES=S.oleracea;
RA Hermann R.G., Alt J., Schiller B., Widger W.R., Cramer W.A.;
RT "Nucleotide sequence of the gene for apocytochrome b-559 on the
RT spinach plastid chromosome: implications for the structure of the
RT membrane protein.";
RL FEBS Lett. 176:239-244(1984).
RN [17]
RP SEQUENCE FROM N.A., AND RNA EDITING OF INITIATOR CODON.
RC SPECIES=C.annuum; STRAIN=cv. Lamuyo; TISSUE=Fruit, and Leaf;
RX MEDLINE=93099270; PubMed=1463853;
RA Kuntz M., Camara B., Weil J.-H., Schantz R.;
RT "The psbI gene from bell pepper (Capsicum annuum): plastid RNA
RT editing also occurs in non-photosynthetic chromoplasts.";
RL Plant Mol. Biol. 20:1185-1188(1992).
RN [18]
RP SEQUENCE FROM N.A.
RC SPECIES=M.cristallinum;
RX MEDLINE=94345017; PubMed=8066140;
RA Porchoefel N.R., Cushman J.C.;
RT "Characterization and expression of photosystem II genes (psbE, psbF,
RT psbL) from the facultative crassulacean acid metabolism plant
RT Mesembryanthemum crystallinum.";
RL Plant Physiol. 105:761-762(1994).
RN [19]
RP SEQUENCE FROM N.A.
RC SPECIES=B.vulgaris; STRAIN=cv. TK81-O; TISSUE=Leaf;
RX MEDLINE=95254673; PubMed=7736615;
RA Kubo T., Yanai Y., Kinoshita T., Mikami T.;
RT "The chloroplast trnT-trnW-petG gene cluster in the mitochondrial
RT genomes of Beta vulgaris, B. trigyna and B. webbiana: evolutionary
RT aspects.";
RL Curr. Genet. 27:285-289(1995).
RN [20]
RP SEQUENCE FROM N.A.
RC SPECIES=P.deltoides; STRAIN=cv. Stoneville D121; TISSUE=Leaf;
RA Naithani S.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [21]
RP SEQUENCE OF 1-15 FROM N.A.
RC SPECIES=T.aestivum, and S.oleracea;
RX MEDLINE=89121082; PubMed=2644131;
RA Ikeuchi M., Takio K., Inoue Y.;
RT "N-terminal sequencing of photosystem II low-molecular-mass proteins.
Query Match 50.0%; Score 22; DB 1; Length 37;
Best Local Similarity 57.1%; Pred. No. 2.3e+02;

```

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LERTEVY 9
Db 12 LNRTSLY 18

RESULT 8
PSBL_CHLEU STANDARD; PRT; 38 AA.
AC P46306;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Chlamydomonas eugametos.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3053;
RN [1]
RP MEDLINE=95269309; PubMed=7750147;
RA Turmel M., Otis C.;
RT "The chloroplast gene cluster containing pabF, pabL, petG and rps3 is
conserved in Chlamydomonas.";
RL Curr. Genet. 27:54-61(1994).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: L29282; AA84157.1; -.
DR PIR; S51366; S51366.
DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4389 MW; DF736DFE7EAC31D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LERTEVY 9
Db 13 LNRTSLY 19

RESULT 9
PSBL_CHLRE STANDARD; PRT; 38 AA.
AC P32974;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Chlamydomonas reinhardtii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LERTEVY 9
Db 13 LNRTSLY 19

RESULT 10
PSBL_CHLUV STANDARD; PRT; 38 AA.
AC P36339;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Chlorella vulgaris.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IAM C-27 / Tamiya;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakauegi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Tsudzuki M.;
RA "Complete nucleotide sequence of the chloroplast genome from the
green alga Chlorella vulgaris: the existence of genes possibly
involved in chloroplast division.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: X66250; CAA46978.1; -.
DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4428 MW; DF73718397EAC31D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LERTEVY 9
Db 13 LNRTSLY 19

```

```

RX MEDLINE=92315354; PubMed=1617741;
RA Fong S.E., Surzycki S.J.;
RT "Organization and structure of plastome pabF, psbl, petG and ORF712
genes in Chlamydomonas reinhardtii.";
RL Curr. Genet. 21:527-530(1992).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: X66250; CAA46978.1; -.
DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4428 MW; DF73718397EAC31D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LERTEVY 9
Db 13 LNRTSLY 19

RESULT 10
PSBL_CHLUV STANDARD; PRT; 38 AA.
AC P36339;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Chlorella vulgaris.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IAM C-27 / Tamiya;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakauegi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Tsudzuki M.;
RA "Complete nucleotide sequence of the chloroplast genome from the
green alga Chlorella vulgaris: the existence of genes possibly
involved in chloroplast division.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AB001684; BAA57901.1; -.
DR PIR; T07254; T07254.
DR InterPro; IPR003372; PSII_Psbl.

```

```

DR Pfam; PF02419; Psbl; 1.
KW ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4388 MW; DF737195DDDEA61D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 13 LNRTSLY 19

RESULT 11
PSBL_EUGR STANDARD; PRT; 38 AA.
AC P12228;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Euglena gracilis.
OC Chloroplast.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z / UTEX 753;
RX MEDLINE=88223485; PubMed=3131027;
RA Cushman J.C., Christopher D.A., Little M.C., Hallick R.B., Price C.A.;
RT "Organization of the psbE, psbF, orf38, and orf42 gene loci on the
RL Euglena gracilis chloroplast genome.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Z / UTEX 753;
RX MEDLINE=93347989; PubMed=8346031;
RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,
RA Orsat B., Spielmann A., Stutz E.;
RT "Complete sequence of Euglena gracilis chloroplast DNA.";
RL Nucleic Acids Res. 21:3537-3544(1993).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
CC B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11874; CAA77914.1; -
DR EMBL; X07073; CAA30110.1; -
DR EMBL; X70810; CAA50097.1; -
DR PIR; S00691; S00691.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4396 MW; DD8E58FD023E90A7 CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 13 LNRTSLY 19

```

```

RESULT 12
PSBL_MARPO STANDARD; PRT; 38 AA.
AC P12185;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Marchantia polymorpha (Liverwort).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;
OC Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=89068687; PubMed=3199436;
RX Fukuzawa H., Kohchi T., Sano T., Shirai H., Umesono K., Inokuchi H.,
RA Ozeki H., Ohshima K.;
RT "Structure and organization of Marchantia polymorpha chloroplast
RT genome. III. Gene organization of the large single copy region from
RT rbcL to trnI(CAU).";
RL J. Mol. Biol. 203:333-351(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,
RA Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,
RA Ozeki H.;
RT "Chloroplast gene organization deduced from complete sequence of
RT liverwort Marchantia polymorpha chloroplast DNA.";
RL Nature 322:572-574(1986).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
CC B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04465; CAA28099.1; -
DR PIR; A05048; A05048.
DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4479 MW; DF73781163D39D6D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 13 LNRTSLY 19

RESULT 13
PSBL_MESVI STANDARD; PRT; 38 AA.
AC Q3MU02;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Mesostigma viride.
OG Chloroplast.

```

OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Otis C., Turmel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
branch of green plant evolution.";
RL Nature 403:649-652(2000).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF166114; AAA43848.1; -
DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4475 MW; C7D0719D78B32B6D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LERTEVY 9
DB 13 LNRTSLY 19

RESULT 14
PSBL ODOSI STANDARD; PRT; 38 AA.
AC P49514;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
Odontella sinensis";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z67753; CAA91712.1; -
DR PIR; S78339; S78339.

DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4365 MW; DF73780E2AF60F6D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LERTEVY 9
DB 13 LNRTSLY 19

RESULT 15
PSBL PINTH STANDARD; PRT; 38 AA.
AC P41617;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Photosystem II reaction center L protein (PSII 5 kDa protein).
GN PSBL.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Pinaceae; Pinus.
OX NCBI_TaxID=3350;
RN [1]
RC STRAIN=NIES-296;
RX MEDLINE=95024047; PubMed=7937893;
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
RA Sugiura M.;
RT "Loss of all ndh genes as determined by sequencing the entire
chloroplast genome of the black pine Pinus thunbergii";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
CC -1- FUNCTION: NOT KNOWN. POSSIBLY A PERIPHERAL COMPONENT OF CYTOCHROME
B559.
CC -1- SIMILARITY: BELONGS TO THE PSBL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D17510; BAA04355.1; -
DR PIR; T07477; T07477.
DR InterPro; IPR003372; PSII_Psbl.
DR Pfam; PF02419; Psbl; 1.
DR ProDom; PD005113; PSII_Psbl; 1.
KW Photosynthesis; Photosystem II; Chloroplast.
SQ SEQUENCE 38 AA; 4505 MW; 23E37813822D8F8D CRC64;

Query Match 50.0%; Score 22; DB 1; Length 38;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LERTEVY 9
DB 13 LNRTSLY 19

Search completed: January 2, 2004, 17:34:15
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 17:29:30 ; Search time 34 Seconds
(without alignments)
68.308 Million cell updates/sec

Title: US-09-641-528B-31040
Perfect score: 44
Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 39761

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	59.1	31	4 Q96FZ1	Q96fz1 homo sapien
2	25	56.8	42	17 Q8ZX76	Q8zx76 pyrobaculum
3	25	56.8	43	16 Q8EZV8	Q8ezv8 leptospira
4	25	56.8	50	11 Q8RI93	Q8ri93 mus musculus
5	25	56.8	50	16 Q8FI43	Q8fi43 escherichia
6	24	54.5	28	6 Q6Z731	Q6z731 canis famil
7	24	54.5	37	16 Q97P03	Q97p03 streptococc
8	24	54.5	38	16 Q8EHS9	Q8ehs9 shewanella
9	24	54.5	40	16 Q97PM6	Q97pm6 streptococc
10	24	54.5	41	16 Q8FI29	Q8fi29 leptospira
11	24	54.5	44	3 Q8J1P2	Q8j1p2 fusarium po
12	24	54.5	44	3 Q8J1F1	Q8j1f1 fusarium sp
13	24	54.5	44	3 Q8J1F0	Q8j1f0 fusarium sp
14	24	54.5	44	3 Q8J1B9	Q8j1b9 fusarium sp
15	24	54.5	44	3 Q8J056	Q8j056 fusarium po
16	24	54.5	44	3 Q8J055	Q8j055 fusarium po

Q8j054 fusarium sp
Q8j053 fusarium ky
Q8n545 homo sapien
Q9hdg0 epichloe fe
Q9etj8 alternaria
Q9etk0 alternaria
Q9hdg1 epichloe ty
Q96ux7 lewia infec
Q96wm6 phaeoacremo
Q9hd18 epichloe cl
Q9etk2 phaeoacremo
Q96ux8 alternaria
Q9c417 xenocylindr
Q9hfz7 epichloe am
Q9hd19 epichloe el
Q9hd19 epichloe ba
Q9etk1 alternaria
Q9hd17 epichloe br
Q9hfz5 epichloe br
Q9etj7 lewia infec
Q96ux6 alternaria
Q9etj9 alternaria
Q9etk3 neotyphodiu
Q96wm7 phaeoacremo
Q8j1b0 epichloe sp
Q8j1a9 neotyphodiu
Q8j1a8 neotyphodiu
Q8j1a7 neotyphodiu
Q8j1a6 neotyphodiu

ALIGNMENTS

RESULT 1

Q96FZ1 ID Q96FZ1 PRELIMINARY; PRT; 31 AA.
AC Q96FZ1; 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010126; AAH0126.1;
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 31 AA; 3577 MW; E47AC48687E6818E CRC64;

Query Match 59.1%; Score 26; DB 4; Length 31;
Best Local Similarity 62.5%; Pred.No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLERTEVY 9
|||
Db 10 TLRRLLEY 17

RESULT 2

Q8ZX76 ID Q8ZX76 PRELIMINARY; PRT; 42 AA.
AC Q8ZX76; 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Hypothetical protein PAE1420.

GN PAE1420.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
ON NCBI_TaxID=13773;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AE009821; AAL63473.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 42 AA; 5055 MW; A8121C27D4984BD0 CRC64;

Query Match 56.8%; Score 25; DB 17; Length 42;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 31 LERGDVY 37

RESULT 3
Q8EZV8 Q8EZV8 PRELIMINARY; PRT; 43 AA.
AC Q8EZV8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA3743.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011530; AAN50941.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 43 AA; 5238 MW; 24A9786E483BF0A4 CRC64;

Query Match 56.8%; Score 25; DB 16; Length 43;
Best Local Similarity 71.4%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 24 LHRTHVY 30

RESULT 4
Q8R193 Q8R193 PRELIMINARY; PRT; 50 AA.
AC Q8R193;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to hypothetical gene supported by AF086316.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025011; AAH25011.1; --
SQ SEQUENCE 50 AA; 6081 MW; C89E2F8B9FE4342 CRC64;

Query Match 56.8%; Score 25; DB 11; Length 50;
Best Local Similarity 83.3%; Pred. No. 6.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATLERT 6
Db 22 ATLQRT 27

RESULT 5
Q8FI43 Q8FI43 PRELIMINARY; PRT; 50 AA.
AC Q8FI43;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN C1605.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=06:HI / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.;
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016759; AAN80073.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 50 AA; 5449 MW; CDB8A5E4FC16694C CRC64;

Query Match 56.8%; Score 25; DB 16; Length 50;
Best Local Similarity 57.1%; Pred. No. 6.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 9 VEATEYI 15

RESULT 6
O62731 O62731 PRELIMINARY; PRT; 28 AA.
AC O62731;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Tropomyosin (Fragment).
GN TPM1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RC SEQUENCE FROM N.A.
RC Liu P.-C., Chen Y.-W., Grob S.E., Katz M.L., Johnson G.S.;
RT "An BagI PCR/RELP in an intron of the canine tropomyosin 1 (TPM1)
gene";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049587; AAC05499.1; --
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00261; Tropomyosin; 1.
FT NON_TER 1


```

FT NON TER 28 28
SQ SEQUENCE 28 AA; 3251 MW; 5694A7F1CA73D30 CRC64;

Query Match
Best Local Similarity 54.5%; Score 24; DB 6; Length 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LERTE 7
Db 1 LERTE 5

RESULT 7
Q97P03 PRELIMINARY; PRT; 37 AA.
AC Q97P03;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Hypothetical protein SP1842.
GN SP1842.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
DR EMBL; AE007476; AAK75915.1; -.
DR TIGR; SP1842; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 37 AA; 4350 MW; EE529568DF2A3F51 CRC64;

Query Match
Best Local Similarity 54.5%; Score 24; DB 16; Length 37;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 19 LEKVELY 25

RESULT 8
Q8EHS9 PRELIMINARY; PRT; 38 AA.
AC Q8EHS9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN S01138.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,

```

```

RA Meyer T., Teapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Mauchtevian J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015558; AAN54208.1; -.
DR TIGR; S01138; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 38 AA; 4524 MW; 9E60C57C452D36EB CRC64;

Query Match
Best Local Similarity 54.5%; Score 24; DB 16; Length 38;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATLERTEVY 9
Db 9 ATLNKQPVY 17

RESULT 9
Q97PM6 PRELIMINARY; PRT; 40 AA.
AC Q97PM6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Hypothetical protein SP1579.
GN SP1579.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
DR EMBL; AE007452; AAK75665.1; -.
DR TIGR; SP1579; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 40 AA; 4721 MW; 9A8CE7B97FC87A29 CRC64;

Query Match
Best Local Similarity 54.5%; Score 24; DB 16; Length 40;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERTEVY 9
Db 26 EKTEVF 31

RESULT 10
Q8F129 PRELIMINARY; PRT; 41 AA.
AC Q8F129;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.

```

GN LA3309.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE011491; AAN50507.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 41 AA; 5086 MW; 6E61C6FDD6F361C6 CRC64;
 Query Match 54.5%; Score 24; DB 16; Length 41;
 Best Local Similarity 50.0%; Pred. No. 8.4e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 TLRETEVY 9
 Db 24 SLHRTDVF 31
 RESULT 11
 Q8J1F2 PRELIMINARY; PRT; 44 AA.
 ID Q8J1F2
 AC Q8J1F2
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Beta-tubulin (Fragment).
 OS Fusarium poae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 OX NCBI_TaxID=36050;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9722-3-1A;
 RA Mach R.L., Kullnig C.M., Kubicek C.P.;
 RA "Beta-tubulin sequence analysis of Fusarium poae, Fusarium pulverosum
 RT and Fusarium sporotrichioides species complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9722-3-1A;
 RA Yli-Mattila T., Konstantinova P., Sletner-Klemsdal S., Mach R.L.,
 RA Kullnig C.M., Kubicek C.P., Bulat S.A., Alekhina I.A.;
 RT "Beta-tubulin IGS and ITS sequence and UP-PCR hybridization analysis
 RT of the Fusarium poae, Fusarium pulverosum and Fusarium
 RT sporotrichioides species complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF404183; AAN76598.1; -
 FT NON_TER 1 1
 FT NON_TER 44 44
 SQ SEQUENCE 44 AA; 4699 MW; CE9B143F287AF2C9 CRC64;
 Query Match 54.5%; Score 24; DB 3; Length 44;
 Best Local Similarity 71.4%; Pred. No. 9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 LERTEVY 9
 Db 32 LERMSVY 38
 RESULT 12
 Q8J1F1 PRELIMINARY; PRT; 44 AA.
 ID Q8J1F1
 AC Q8J1F1
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Beta-tubulin (Fragment).
 OS Fusarium sporotrichioides.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 OX NCBI_TaxID=5514;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=93-453;
 RA Mach R.L., Kullnig C.M., Kubicek C.P.;
 RA "Beta-tubulin sequence analysis of Fusarium poae, Fusarium pulverosum
 RT and Fusarium sporotrichioides species complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=93-453;
 RA Yli-Mattila T., Konstantinova P., Sletner-Klemsdal S., Mach R.L.,
 RA Kullnig C.M., Kubicek C.P., Bulat S.A., Alekhina I.A.;
 RT "Beta-tubulin IGS and ITS sequence and UP-PCR hybridization analysis
 RT of the Fusarium poae, Fusarium pulverosum and Fusarium
 RT sporotrichioides species complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF404185; AAN76600.1; -
 FT NON_TER 1 1
 FT NON_TER 44 44
 SQ SEQUENCE 44 AA; 4828 MW; CF9E0FC459AAF2C9 CRC64;
 Query Match 54.5%; Score 24; DB 3; Length 44;
 Best Local Similarity 71.4%; Pred. No. 9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 LERTEVY 9
 Db 32 LERMSVY 38
 RESULT 13
 Q8J1F0 PRELIMINARY; PRT; 44 AA.
 ID Q8J1F0
 AC Q8J1F0
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Beta-tubulin (Fragment).
 OS Fusarium sporotrichioides.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 OX NCBI_TaxID=5514;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IBT 8804;
 RA Mach R.L., Kullnig C.M., Kubicek C.P.;
 RA "Beta-tubulin sequence analysis of Fusarium poae, Fusarium pulverosum
 RT and Fusarium sporotrichioides species complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IBT 8804;
 RA Yli-Mattila T., Konstantinova P., Sletner-Klemsdal S., Mach R.L.,
 RA Kullnig C.M., Kubicek C.P., Bulat S.A., Alekhina I.A.;
 RT "Beta-tubulin IGS and ITS sequence and UP-PCR hybridization analysis
 RT of the Fusarium poae, Fusarium pulverosum and Fusarium
 RT sporotrichioides species complex.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF404187; AAN76602.1; -
 FT NON_TER 1 1
 FT NON_TER 44 44
 SQ SEQUENCE 44 AA; 4744 MW; CE9E0E2A575AF2C9 CRC64;
 Query Match 54.5%; Score 24; DB 3; Length 44;
 Best Local Similarity 71.4%; Pred. No. 9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 LERTEVY 9
 Db 32 LERMSVY 38

```

RESULT 14
Q8J1E9 PRELIMINARY; PRT; 44 AA.
AC Q8J1E9
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Beta-tubulin (Fragment).
OS Fusarium sporotrichioides.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Fusarium.
OX NCBI_TaxID=5514;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-92;
RA Mach R.L., Kullnig C.M., Kubicek C.P.;
RT "Beta-tubulin sequence analysis of Fusarium poae, Fusarium pulverosum
RT and Fusarium sporotrichioides species complex.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=97-92;
RA Yli-Mattila T., Konstantinova P., Sletner-Klemsdal S., Mach R.L.,
RA Kullnig C.M., Kubicek C.P., Bulat S.A., Alekhina I.A.;
RT "Beta-tubulin IGS and ITS sequence and up-PCR hybridization analysis
RT of the Fusarium poae, Fusarium pulverosum and Fusarium
RT sporotrichioides species complex.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF404190; AAN76605.1; -
FT NON_TER 1
FT NON_TER 44
SQ SEQUENCE 44 AA; 4745 MW; CE9E056E3AAF2C9 CRC64;

Query Match 54.5%; Score 24; DB 3; Length 44;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 32 LERMSVY 38

RESULT 15
Q8J056 PRELIMINARY; PRT; 44 AA.
AC Q8J056
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Beta-tubulin (Fragment).
OS Fusarium poae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Fusarium.
OX NCBI_TaxID=36050;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Various strains;
RA Mach R.L., Kullnig C.M., Kubicek C.P.;
RT "Beta-tubulin sequence analysis of Fusarium poae, Fusarium pulverosum
RT and Fusarium sporotrichioides species complex.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Various strains;
RA Yli-Mattila T., Konstantinova P., Sletner-Klemsdal S., Mach R.L.,
RA Kullnig C.M., Kubicek C.P., Bulat S.A., Alekhina I.A.;
RT "Beta-tubulin IGS and ITS sequence and up-PCR hybridization analysis
RT of the Fusarium poae, Fusarium pulverosum and Fusarium
RT sporotrichioides species complex.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF404184; AAN76599.1; -

```

```

DR EMBL; AF404192; AAN76607.1; -
DR EMBL; AF404193; AAN76608.1; -
DR EMBL; AF404194; AAN76609.1; -
DR EMBL; AF404195; AAN76610.1; -
DR EMBL; AF404196; AAN76611.1; -
DR EMBL; AF404197; AAN76612.1; -
DR EMBL; AF404198; AAN76613.1; -
DR EMBL; AF404199; AAN76614.1; -
DR EMBL; AF404200; AAN76615.1; -
DR EMBL; AF404201; AAN76616.1; -
DR EMBL; AF404202; AAN76617.1; -
DR EMBL; AF404203; AAN76618.1; -
DR EMBL; AF404204; AAN76619.1; -
DR EMBL; AF404205; AAN76620.1; -
DR EMBL; AF404206; AAN76621.1; -
DR EMBL; AF404207; AAN76622.1; -
DR EMBL; AF404208; AAN76623.1; -
DR EMBL; AF404209; AAN76624.1; -
DR EMBL; AF404210; AAN76625.1; -
DR EMBL; AF404211; AAN76626.1; -
DR EMBL; AF404212; AAN76627.1; -
DR EMBL; AF404213; AAN76628.1; -
DR EMBL; AF404214; AAN76629.1; -
DR EMBL; AF404215; AAN76630.1; -
DR EMBL; AF404216; AAN76631.1; -
DR EMBL; AF404217; AAN76632.1; -
DR EMBL; AF404218; AAN76633.1; -
DR EMBL; AF404219; AAN76634.1; -
DR EMBL; AF404220; AAN76635.1; -
DR EMBL; AF404221; AAN76636.1; -
DR EMBL; AF404222; AAN76637.1; -
DR EMBL; AF404223; AAN76638.1; -
DR EMBL; AF404224; AAN76639.1; -
DR EMBL; AF404225; AAN76640.1; -
DR EMBL; AF404226; AAN76641.1; -
DR EMBL; AF404227; AAN76642.1; -
DR EMBL; AF404228; AAN76643.1; -
DR EMBL; AF404229; AAN76644.1; -
DR EMBL; AF404230; AAN76645.1; -
DR EMBL; AF404231; AAN76646.1; -
DR EMBL; AF404232; AAN76647.1; -
DR EMBL; AF404233; AAN76648.1; -
DR EMBL; AF404234; AAN76649.1; -
DR EMBL; AF404235; AAN76650.1; -
DR EMBL; AF404236; AAN76651.1; -
DR EMBL; AF404237; AAN76652.1; -
FT NON_TER 1
FT NON_TER 44
SQ SEQUENCE 44 AA; 4665 MW; CB9B1B9F827AF2C9 CRC64;

Query Match 54.5%; Score 24; DB 3; Length 44;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LERTEVY 9
Db 32 LERMSVY 38

```

Search completed: January 2, 2004, 17:35:02
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 17:30:50 ; Search time 22 Seconds
(without alignments)
17.309 Million cell updates/sec

Title: US-09-641-528B-31040

Perfect score: 44

Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 206161

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/iaa/5A COMB.pdp.*
- 2: /cgn2_6/prodata/iaa/5B COMB.pdp.*
- 3: /cgn2_6/prodata/iaa/6A COMB.pdp.*
- 4: /cgn2_6/prodata/iaa/6B COMB.pdp.*
- 5: /cgn2_6/prodata/iaa/PCTUS COMB.pdp.*
- 6: /cgn2_6/prodata/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	61.4	12	3	US-09-439-897-54
2	27	61.4	35	3	US-09-439-897-64
3	26	59.1	25	1	US-07-941-651-6
4	26	59.1	25	1	US-08-279-996-6
5	26	59.1	25	6	5248606-44
6	26	59.1	29	4	US-09-354-231B-31
7	26	59.1	29	4	US-09-354-231B-33
8	26	59.1	29	4	US-09-354-231B-34
9	26	59.1	29	4	US-09-128-602B-31
10	26	59.1	29	4	US-09-128-602B-33
11	26	59.1	29	4	US-09-128-602B-34
12	25	56.8	18	6	517197-14
13	24	54.5	8	4	US-09-367-777-105
14	24	54.5	9	4	US-09-400-564-5
15	24	54.5	12	4	US-09-400-564-3
16	24	54.5	13	4	US-09-400-564-2
17	24	54.5	14	4	US-09-400-564-1
18	24	54.5	15	3	US-08-159-339A-1174
19	24	54.5	15	4	US-09-400-564-4
20	24	54.5	27	4	US-09-400-564-13
21	24	54.5	36	3	US-09-082-279B-871
22	24	54.5	36	4	US-09-315-304B-871
23	24	54.5	36	4	US-09-834-784-871
24	24	54.5	44	2	US-08-691-814B-46
25	24	54.5	48	4	US-09-314-268-95
26	23	52.3	16	2	US-08-484-905-27
27	23	52.3	16	3	US-08-481-985B-27

28	23	52.3	16	3	US-08-370-476-27	Sequence 27, Appl
29	23	52.3	21	4	US-09-443-339A-6	Sequence 6, Appl
30	23	52.3	27	4	US-09-245-712-12	Sequence 12, Appl
31	23	52.3	27	4	US-09-245-712-13	Sequence 13, Appl
32	23	52.3	27	4	US-09-245-712-14	Sequence 14, Appl
33	23	52.3	29	4	US-09-354-231B-32	Sequence 32, Appl
34	23	52.3	29	4	US-09-128-602B-32	Sequence 32, Appl
35	23	52.3	38	1	US-07-977-630-67	Sequence 67, Appl
36	23	52.3	47	1	US-08-582-257-14	Sequence 14, Appl
37	23	52.3	47	2	US-08-582-298-14	Sequence 14, Appl
38	22	50.0	10	1	US-08-347-198A-13	Sequence 13, Appl
39	22	50.0	10	3	US-08-335-844A-37	Sequence 37, Appl
40	22	50.0	10	3	US-08-335-844A-52	Sequence 52, Appl
41	22	50.0	10	4	US-09-129-366-37	Sequence 37, Appl
42	22	50.0	10	4	US-09-129-366-52	Sequence 52, Appl
43	22	50.0	20	1	US-08-273-669-7	Sequence 7, Appl
44	22	50.0	20	2	US-08-954-724-7	Sequence 7, Appl
45	22	50.0	20	2	US-08-965-947-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-439-897-54
; Sequence 54, Application US/09439897
; Patent No. 6277558
; GENERAL INFORMATION:
; APPLICANT: Hudson, Billy G
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides
; FILE REFERENCE: 95-1263-C
; CURRENT APPLICATION NUMBER: US/09/439,897
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: construct C3 alpha1
US-09-439-897-54

Query Match 61.4%; Score 27; DB 3; Length 12;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEVY 9
|:|:|:
Db 1 TIERSEMP 8

RESULT 2
US-09-439-897-64
; Sequence 64, Application US/09439897
; Patent No. 6277558
; GENERAL INFORMATION:
; APPLICANT: Hudson, Billy G
; TITLE OF INVENTION: Alpha-3 Chain Type IV Collagen Polypeptides
; FILE REFERENCE: 95-1263-C
; CURRENT APPLICATION NUMBER: US/09/439,897
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: construct C8 alpha1
US-09-439-897-64

Query Match 61.4%; Score 27; DB 3; Length 35;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TLRTVEVY 9
|:|:|:|:
Db 1 TIERSEMF 8

RESULT 3
US-07-941-651-6
; Sequence 6, Application US/07941651
; Patent No. 5332808
; GENERAL INFORMATION:
; APPLICANT: Boston, Rebecca S.
; APPLICANT: Bass, Henry W.
; APPLICANT: O'Brian, Gregory R.
; TITLE OF INVENTION: DNA Encoding A Ribosome Inactivating
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and
; ADDRESSEE: Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5332808th Carolina
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/941,651
; FILING DATE: 19920908
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5051-199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-941-651-6

Query Match 59.1%; Score 26; DB 1; Length 25;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATLRETEV 8
|:|:|:|:
Db 2 ATLEEEV 9

RESULT 4
US-08-279-996-6
; Sequence 6, Application US/08279996
; Patent No. 5552140
; GENERAL INFORMATION:
; APPLICANT: Boston, Rebecca S.
; APPLICANT: Bass, Henry W.
; APPLICANT: O'Brian, Gregory R.
; TITLE OF INVENTION: DNA Encoding A Ribosome Inactivating
; TITLE OF INVENTION: Protein

; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and
; ADDRESSEE: Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5552140th Carolina
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,996
; FILING DATE: 25-JUL-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/941,651
; FILING DATE: 08-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5051-199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-279-996-6

Query Match 59.1%; Score 26; DB 1; Length 25;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATLRETEV 8
|:|:|:|:
Db 2 ATLEEEV 9

RESULT 5
5248606-44
; Patent No. 5248606
; APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,
; ALICE E.R.
; TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND
; ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATING
; NUMBER OF SEQUENCES: 49
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/535,636
; FILING DATE: 11-JUN-1990
; SEQ ID NO: 44:
; LENGTH: 25
5248606-44

Query Match 59.1%; Score 26; DB 6; Length 25;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATLRETEV 8
|:|:|:|:
Db 2 ATLEEEV 9

RESULT 6
US-09-354-231B-31
; Sequence 31, Application US/09354231B

```
/ Patent No. 6342658
/ GENERAL INFORMATION:
/ APPLICANT: DeBonte, Lorin R.
/ APPLICANT: Shortosh, Basil S.
/ TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
/ FILE REFERENCE: 07148-063002
/ CURRENT APPLICATION NUMBER: US/09/354,231B
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: US 08/874,109
/ PRIOR FILING DATE: 1997-06-12
/ NUMBER OF SEQ ID NOS: 69
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 31
/ LENGTH: 29
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
US-09-354-231B-31

Query Match
Best Local Similarity 59.1%; Score 26; DB 4; Length 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLTEREVY 9
DB 21 SLRDEVF 28

RESULT 7
US-09-354-231B-33
/ Sequence 33, Application US/09354231B
/ Patent No. 6342658
/ GENERAL INFORMATION:
/ APPLICANT: DeBonte, Lorin R.
/ APPLICANT: Shortosh, Basil S.
/ TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
/ FILE REFERENCE: 07148-063002
/ CURRENT APPLICATION NUMBER: US/09/354,231B
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: US 08/874,109
/ PRIOR FILING DATE: 1997-06-12
/ NUMBER OF SEQ ID NOS: 69
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 33
/ LENGTH: 29
/ TYPE: PRT
/ ORGANISM: Zea mays
US-09-354-231B-33

Query Match
Best Local Similarity 59.1%; Score 26; DB 4; Length 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLTEREVY 9
DB 21 SLRDEVF 28

RESULT 8
US-09-354-231B-34
/ Sequence 34, Application US/09354231B
/ Patent No. 6342658
/ GENERAL INFORMATION:
/ APPLICANT: DeBonte, Lorin R.
/ APPLICANT: Shortosh, Basil S.
/ TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
/ FILE REFERENCE: 07148-063002
/ CURRENT APPLICATION NUMBER: US/09/354,231B
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: US 08/874,109
/ PRIOR FILING DATE: 1997-06-12
/ NUMBER OF SEQ ID NOS: 69
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 34
```

```
/ LENGTH: 29
/ TYPE: PRT
/ ORGANISM: Ricinus communis
US-09-354-231B-34

Query Match
Best Local Similarity 59.1%; Score 26; DB 4; Length 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLTEREVY 9
DB 21 SLRDEVF 28

RESULT 9
US-09-128-602B-31
/ Sequence 31, Application US/09128602B
/ Patent No. 6414223
/ GENERAL INFORMATION:
/ APPLICANT: Kodali, Dharmma
/ APPLICANT: Pan, Zhegong
/ APPLICANT: DeBonte, Lorin R.
/ TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATED
/ FILE REFERENCE: 07148-072001
/ CURRENT APPLICATION NUMBER: US/09/128,602B
/ CURRENT FILING DATE: 1998-08-03
/ NUMBER OF SEQ ID NOS: 68
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 31
/ LENGTH: 29
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
US-09-128-602B-31

Query Match
Best Local Similarity 59.1%; Score 26; DB 4; Length 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLTEREVY 9
DB 21 SLRDEVF 28

RESULT 10
US-09-128-602B-33
/ Sequence 33, Application US/09128602B
/ Patent No. 6414223
/ GENERAL INFORMATION:
/ APPLICANT: Kodali, Dharmma
/ APPLICANT: Pan, Zhegong
/ APPLICANT: DeBonte, Lorin R.
/ TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATED
/ FILE REFERENCE: 07148-072001
/ CURRENT APPLICATION NUMBER: US/09/128,602B
/ CURRENT FILING DATE: 1998-08-03
/ NUMBER OF SEQ ID NOS: 68
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 33
/ LENGTH: 29
/ TYPE: PRT
/ ORGANISM: Zea mays
US-09-128-602B-33

Query Match
Best Local Similarity 59.1%; Score 26; DB 4; Length 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLTEREVY 9
DB 21 SLRDEVF 28
```

RESULT 11
US-09-128-602B-34
; Sequence 34, Application US/09128602B
; Patent No. 6414223
; GENERAL INFORMATION:
; APPLICANT: Kodali, Dharna
; APPLICANT: Pan, Zhegong
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATED
; FILE REFERENCE: 07148-072001
; CURRENT APPLICATION NUMBER: US/09/128,602B
; CURRENT FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Ricinus communis
US-09-128-602B-34

Query Match 59.1%; Score 26; DB 4; Length 29;
Best Local Similarity 62.5%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRETVY 9
Db 21 SLERDEVF 28

RESULT 12
5177197-14
; Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
; LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 14:
; LENGTH: 18
5177197-14

Query Match 56.8%; Score 25; DB 6; Length 18;
Best Local Similarity 57.1%; Pred. No. 59;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LRETVY 9
Db 11 IEETDVY 17

RESULT 13
US-09-367-777-105
; Sequence 105, Application US/09367777
; Patent No. 6562598
; GENERAL INFORMATION:
; APPLICANT: Himmelapach, Michele
; Pfleiderer, Michael
; Falkner, Falko-Gunter
; Bibl, Johann
; Dorner, Friedrich
; Schlokot, Uwe
; TITLE OF INVENTION: Factor X Deletion Mutants
; and Analogues Thereof
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/367,777
FILING DATE: 10-NOV-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AT A 336/97
FILING DATE: 27-FEB-1997
APPLICATION NUMBER: WO PCT/AT98/00046
FILING DATE: 27-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Auserhus, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 20695D-0009000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-09-367-777-105

Query Match 54.5%; Score 24; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLRT 6
Db 1 TLRT 5

RESULT 14
US-09-400-564-5
; Sequence 5, Application US/09400564
; Patent No. 6350574
; GENERAL INFORMATION:
; APPLICANT: Montelaro, Ronald C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Jolley, Michael E.
; APPLICANT: Nasir, Mohammad S.
; TITLE OF INVENTION: A Fluorescence Polarization-Based Diagnostic Assay
; FILE REFERENCE: Case No. 6350574 99,579
; CURRENT APPLICATION NUMBER: US/09/400,564
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 60/101,553
; NUMBER OF SEQ ID NOS: 21
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Equine infectious anemia virus
US-09-400-564-5

Query Match 54.5%; Score 24; DB 4; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.5e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LRETVY 9
Db 11 IEETDVY 17

Db 2 IERTV 8

RESULT 15
US-09-400-564-3
; Sequence 3, Application US/09400564
; Patent No. 6350574
; GENERAL INFORMATION:
; APPLICANT: Montelaro, Ronald C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Jolley, Michael E.
; APPLICANT: Nasir, Mohammad S.
; TITLE OF INVENTION: A Fluorescence Polarization-Based Diagnostic Assay
; TITLE OF INVENTION: for Equine Infectious Anemia Virus
; FILE REFERENCE: Case No. 6350574 99,579
; CURRENT APPLICATION NUMBER: US/09/400,564
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 60/101,553
; NUMBER OF SEQ ID NOS: 21
; SEQ ID NO 3
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Equine infectious anemia virus
US-09-400-564-3

Query Match 54.5%; Score 24; DB 4; Length 12;
Best Local Similarity 57.1%; Pred. No. 60;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LERTV 9

Db 2 IERTV 8

Search completed: January 2, 2004, 17:35:36
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 17:35:06 ; Search time 31 Seconds
(without alignments)
58.251 Million cell updates/sec

Title: US-09-641-528B-31040
Perfect score: 44
Sequence: 1 ATLERTEVY 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 216031

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	70.5	21	15	US-10-206-699-245
2	29	65.9	9	8	US-08-821-739A-38
3	28	63.6	25	11	US-09-974-879-368
4	28	63.6	25	11	US-09-305-736-368
5	28	63.6	25	12	US-09-818-683-368
6	27	61.4	15	15	US-10-206-699-50
7	27	61.4	35	12	US-10-074-511-91
8	26	59.1	29	9	US-09-995-297-31
9	26	59.1	29	9	US-09-995-297-33
10	26	59.1	29	9	US-09-995-297-34
11	26	59.1	29	12	US-09-771-904-31
12	26	59.1	29	12	US-09-771-904-33
13	26	59.1	29	12	US-09-771-904-34
14	26	59.1	49	9	US-09-864-761-38416
15	25	56.8	29	9	US-09-864-761-35300

16	24	54.5	8	12	US-10-348-504-105
17	24	54.5	36	9	US-09-864-761-35959
18	24	54.5	36	12	US-10-351-641-871
19	24	54.5	44	9	US-09-764-869-894
20	24	54.5	44	11	US-09-774-639-356
21	24	54.5	44	11	US-09-969-730-210
22	24	54.5	44	15	US-10-091-504-894
23	24	54.5	48	9	US-09-864-761-47225
24	24	54.5	48	12	US-10-008-524A-95
25	24	54.5	48	12	US-10-350-719-95
26	24	54.5	49	9	US-09-864-761-47644
27	23	52.3	9	12	US-09-965-738-155
28	23	52.3	16	15	US-10-225-567A-1214
29	23	52.3	18	12	US-10-116-275-333
30	23	52.3	18	15	US-10-059-271-17
31	23	52.3	21	9	US-09-864-761-41221
32	23	52.3	23	9	US-09-864-675-16
33	23	52.3	23	11	US-09-895-298-188
34	23	52.3	24	9	US-09-864-761-35834
35	23	52.3	28	12	US-10-195-730-343
36	23	52.3	29	9	US-09-995-297-32
37	23	52.3	29	12	US-09-771-904-32
38	23	52.3	29	12	US-10-339-740-249
39	23	52.3	30	9	US-09-864-761-40825
40	23	52.3	31	9	US-09-864-761-46822
41	23	52.3	32	15	US-10-174-410-221
42	23	52.3	36	9	US-09-864-761-48948
43	23	52.3	36	15	US-10-011-585A-175
44	23	52.3	37	12	US-10-139-794-61
45	23	52.3	39	9	US-09-864-761-37976

ALIGNMENTS

RESULT 1

US-10-206-699-245
; Sequence 245, Application US/10206699
; Publication No. US20030100510A1
; GENERAL INFORMATION:
; APPLICANT: Sundaramoorthy, M.
; APPLICANT: Hudson, B.
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NC1 Domain Hexamer
; FILE REFERENCE: MBHB 01-1017
; CURRENT APPLICATION NUMBER: US/10/206,699
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 245
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-206-699-245

Query Match 70.5%; Score 31; DB 15; Length 21;
Best Local Similarity 55.6%; Pred. No. 6.5;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATLERTEVY 9

Db 4 ATIERSEMF 12

RESULT 2

US-08-821-739A-38

; Sequence 38, Application US/08821739A
; Publication No. US20020168374A1
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Eteban
; TITLE OF INVENTION: HLA Binding Peptides and Their Uses
; FILE REFERENCE: 2060.005000A
; CURRENT APPLICATION NUMBER: US/08/821,739A
; CURRENT FILING DATE: 1999-03-20
; PRIOR APPLICATION NUMBER: 60/013,833
; PRIOR FILING DATE: 1996-03-21
; PRIOR APPLICATION NUMBER: 08/589,107
; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 08/451,913
; PRIOR FILING DATE: 1995-05-26
; PRIOR APPLICATION NUMBER: 08/347,610
; PRIOR FILING DATE: 1994-12-01
; PRIOR APPLICATION NUMBER: 08/186,266
; PRIOR FILING DATE: 1994-01-25
; PRIOR APPLICATION NUMBER: 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: 08/027,746
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: 07/926,666
; PRIOR FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-821-739A-38

Query Match 65.9%; Score 29; DB 8; Length 9;
Best Local Similarity 71.4%; Pred. No. 6.5e+05;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LERTEVY 9
:|||||
Db 2 VERTDVY 8

RESULT 3
US-09-974-879-368
; Sequence 368, Application US/09974879
; Publication No. US20030028003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818,683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305,736
; PRIOR FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,988

; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 368
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-974-879-368

Query Match 63.6%; Score 28; DB 11; Length 25;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LERTEVY 9
:|||||
Db 8 LERTEVY 14

RESULT 4
US-09-305-736-368
; Sequence 368, Application US/09305736
; Publication No. US2003008078A1
; GENERAL INFORMATION:
; APPLICANT: Feng et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P1
; CURRENT APPLICATION NUMBER: US/09/305,736
; CURRENT FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/064,985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/066,094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066,095

; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,090
; EARLIER FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 612
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 368
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-305-736-368

Query Match 63.6%; Score 28; DB 11; Length 25;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LERTEVY 9
|||
Db 8 LELTEVY 14

RESULT 5
US-09-818-683-368
; Sequence 368, Application US/09818683
; Publication No. US20030211472A1
; GENERAL INFORMATION:
; APPLICANT: Feng et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P1
; CURRENT APPLICATION NUMBER: US/09/818,683
; CURRENT FILING DATE: 2001-03-28
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 612
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 368
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-818-683-368

Query Match 63.6%; Score 28; DB 12; Length 25;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LERTEVY 9
|||
Db 8 LELTEVY 14

RESULT 6
US-10-206-699-50
; Sequence 50, Application US/10206699
; Publication No. US20030100510A1
; GENERAL INFORMATION:
; APPLICANT: Sundaramoorthy, M.
; APPLICANT: Hudson, B.
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NC1 Domain Hexamer
; FILE REFERENCE: MBHB 01-1017
; CURRENT APPLICATION NUMBER: US/10/206,699
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-206-699-50

Query Match 61.4%; Score 27; DB 15; Length 15;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLERTEVY 9
|:|:|:|:
Db 1 TIERSEMF 8

RESULT 7
US-10-074-511-91
; Sequence 91, Application US/10074511
; Publication No. US20030176672A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karta, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Liu, Chenghua
; APPLICANT: Sun, Yongming
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pro
; FILE REFERENCE: DEX-0314
; CURRENT APPLICATION NUMBER: US/10/074,511
; CURRENT FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/268,289
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-074-511-91

Query Match 61.4%; Score 27; DB 12; Length 35;
Best Local Similarity 71.4%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLERTEV 8
|||
Db 22 TLRSTEI 28

RESULT 8
US-09-995-297-31
; Sequence 31, Application US/09995297
; Patent No. US20020092038A1
; GENERAL INFORMATION:
; APPLICANT: Kodali, Dharna
; APPLICANT: Fan, Zhegong
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATED
; FILE REFERENCE: 07148-072002
; CURRENT APPLICATION NUMBER: US/09/995,297
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: US 09/128,602
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-995-297-31

Query Match 59.1%; Score 26; DB 9; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRTEVY 9
:|||||:
Db 21 SLRDEVF 28

RESULT 9

US-09-995-297-33
; Sequence 33, Application US/09995297
; Patent No. US20020092038A1
; GENERAL INFORMATION:
; APPLICANT: Kodali, Dharna
; APPLICANT: Fan, Zhegong
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATED
; FILE REFERENCE: 07148-072002
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: US/09/995,297
; PRIOR FILING DATE: 2001-11-27
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Zea mays
US-09-995-297-33

Query Match 59.1%; Score 26; DB 9; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRTEVY 9
:|||||:
Db 21 SLRDEVF 28

RESULT 10

US-09-995-297-34
; Sequence 34, Application US/09995297
; Patent No. US20020092038A1
; GENERAL INFORMATION:
; APPLICANT: Kodali, Dharna
; APPLICANT: Fan, Zhegong
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS, SEEDS AND OILS HAVING AN ELEVATED TOTAL MONOUNSATURATED
; FILE REFERENCE: 07148-072002
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: US/09/995,297
; PRIOR FILING DATE: 2001-11-27
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Ricinus communis
US-09-995-297-34

Query Match 59.1%; Score 26; DB 9; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRTEVY 9
:|||||:
Db 21 SLRDEVF 28

RESULT 11

US-09-771-904-31
; Sequence 31, Application US/09771904

; Publication No. US20030131379A1
; GENERAL INFORMATION:
; APPLICANT: DeBonte, Lorin R.
; APPLICANT: Fan, Zhegong
; APPLICANT: Miao, Guo-Hua
; TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
; FILE REFERENCE: 07148-063003
; CURRENT APPLICATION NUMBER: US/09/771,904
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 08/874,109
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-771-904-31

Query Match 59.1%; Score 26; DB 12; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRTEVY 9
:|||||:
Db 21 SLRDEVF 28

RESULT 12

US-09-771-904-33
; Sequence 33, Application US/09771904
; Publication No. US20030131379A1
; GENERAL INFORMATION:
; APPLICANT: DeBonte, Lorin R.
; APPLICANT: Fan, Zhegong
; APPLICANT: Miao, Guo-Hua
; TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
; FILE REFERENCE: 07148-063003
; CURRENT APPLICATION NUMBER: US/09/771,904
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 08/874,109
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Zea mays
US-09-771-904-33

Query Match 59.1%; Score 26; DB 12; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRTEVY 9
:|||||:
Db 21 SLRDEVF 28

RESULT 13

US-09-771-904-34
; Sequence 34, Application US/09771904
; Publication No. US20030131379A1
; GENERAL INFORMATION:
; APPLICANT: DeBonte, Lorin R.
; APPLICANT: Fan, Zhegong
; APPLICANT: Miao, Guo-Hua
; TITLE OF INVENTION: FATTY ACID DESATURASES AND MUTANT SEQUENCES THEREOF
; FILE REFERENCE: 07148-063003
; CURRENT APPLICATION NUMBER: US/09/771,904
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 08/874,109
; PRIOR FILING DATE: 1997-06-12

; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Ricinus communis
US-09-771-904-34

Query Match 59.1%; Score 26; DB 12; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TLRTVEY 9
:|||||:
Db 21 SLRDEVP 28

RESULT 14
US-09-864-761-38416
; Sequence 38416, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38416
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: MAP TO AL035701.13
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EST HUMAN HIT: AW270089.1, EVALUE 1.00e-15
; OTHER INFORMATION: SWISSPROT HIT: P22413, EVALUE 4.60e-02
US-09-864-761-38416

Query Match 59.1%; Score 26; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RTEVY 9
:|||||:
Db 1 RTEVY 5

RESULT 15
US-09-864-761-35300
; Sequence 35300, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35300
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL17378.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AV752481.1, EVALUE 7.00e-11
; OTHER INFORMATION: SWISSPROT HIT: P22413, EVALUE 5.00e-12
US-09-864-761-35300

```

```

Query Match          56.8%; Score 25; DB 9; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      2 TLRTVEY 9
      | | | | |
Db     18 TFENIEY 25

```

```

Search completed: January 2, 2004, 17:40:27
Job time : 32 secs

```